Theory of Stochastic Processes
Vol. 16 (32), no. 2, 2010, pp. 106–119

ALEXANDER MARYNYCH

ON THE ASYMPTOTICS OF MOMENTS OF LINEAR RANDOM RECURRENCES

We propose a new method of analyzing the asymptotics of moments of certain linear random recurrences which is based on the technique of iterative functions. By using the method, we show that the moments of the number of collisions and the absorption time in the Poisson–Dirichlet coalescent behave like the powers of the "log star" function which grows slower than any iteration of the logarithm, and thereby we prove a weak law of large numbers. Finally, we discuss merits and limitations of the method and give several examples related to beta coalescents, recursive algorithms, and random trees.

1. Introduction and main result

A *linear random recurrence* is a sequence of random variables $\{X_n, n \in \mathbb{N}\}$ which satisfies the distributional equality

$$X_1 = 0, \quad X_n \overset{d}{=} V_n + \sum_{r=1}^{K} A_r(n) X^{(r)}_{I^n_{(r)}}, \quad n \geq 2, \tag{1}$$

where X_n is some parameter of a problem of size n which splits into $K \geq 1$ subproblems of random sizes $I^n_{(r)} \in \{1, \dots, n\}$. For every $r = 1, \dots, K$, the sequence $\{X^{(r)}_k, k \in \mathbb{N}\}$ which corresponds to the contribution of subgroup r is a distributional copy of $\{X_k, k \in \mathbb{N}\}$, V_n is a random toll term, and $A_r(n) > 0$ is a random weight of subgroup r. It is assumed that $\{(I^n_{(1)}, \dots, I^n_{(K)}, A_1(n), \dots, A_K(n), V_n), n \geq 2\}$, $\{X^{(1)}_n, n \in \mathbb{N}\}, \dots, \{X^{(K)}_n, n \in \mathbb{N}\}$ are independent.

Random recurrences (1), often in a simplified form with $K = 1$, arise in various areas of applied probability such as random regenerative structures [9, 11], random trees [5, 7, 25, 26], coalescent theory [6, 10, 12, 16], absorption times in non-increasing Markov chains [13, 2], recursive algorithms [15, 24, 27, 28], random walks with barrier [17, 18], to name but a few.

The first step of the asymptotic analysis of recurrences (1) is to find the asymptotics of moments $\mathbb{E}X_n^k$ and central moments $\mathbb{E}(X_n - \mathbb{E}X_n)^k$, as $n \to \infty$. This problem reduces to studying the recurrent equations of the form

$$a_1 = 0, \;\; a_n = b_n + \sum_{k=1}^{n-1} c_{nk} a_k, \;\; n \geq 2, \tag{2}$$

where $\{b_n, n \in \mathbb{N}\}$ and $\{c_{nk}, n \in \mathbb{N}, k < n\}$ are given numerical sequences. The purpose of the present paper is to propose a new method of obtaining the first-order asymptotics of solutions to (2), as $n \to \infty$.

Although the asymptotic analysis of recurrences (2) is a hard analytic problem, some more or less efficient methods have been elaborated to date. Evidently, the most popular existing approach is the *method of singular analysis of generating functions* [5, 8]. The method gives a very precise information on the asymptotic behavior, whenever there is a

---

2000 *Mathematics Subject Classification.* Primary 60F05; Secondary 60C05.
Key words and phrases. moments, Poisson–Dirichlet coalescent, linear recurrence.

tractable functional relation between the generating functions of the sequences involved. The idea of a *repertoire method* proposed in [15] can be briefly described as follows. First, we build up a repertoire $\{b_n^{(\alpha)},\ \alpha \in A\}$, where A is a finite set, of inhomogeneous terms of (2) by choosing sequences $\{a_n^{(\alpha)},\ n \in \mathbb{N}\}$ such that the sum in (2) is tractable. Then we construct the solution a_n to (2) with the inhomogeneous term b_n as a linear combination of solutions $a_n^{(\alpha)},\ \alpha \in A$. Finally, we mention a method proposed in [1] and further developed in [28] which is based on the harmonic analysis and potential theory.

Initially, our method which is based on the technique of iterative functions was developed in order to find the asymptotics of moments of the number of collisions X_n and the absorption time T_n in the Poisson–Dirichlet coalescent (this problem was raised by Martin Möhle in the early 2008). Below, we prove that both $\mathbb{E}X_n^k$ and $\mathbb{E}T_n^k,\ k \in \mathbb{N}$, behave like the powers of a "log star" function which grows slower than any iteration of the logarithm[1]. This somewhat exotic behavior of the moments partially explains the fact that we have not been able to apply either of previously known (to us) methods to tackle the problem. To our knowledge, the "log star" asymptotics arises not often. In particular, we are only aware of two applied models which exhibit such a behavior: (a) the number of distinguishable alleles according to the Ohta–Kimura model of neutral mutation [20], and (b) the average complexity of Delaunay triangulation of the Euclidean minimum spanning tree [4]. The number of collisions and the absorption time in the Poisson–Dirichlet coalescent are interesting, yet particular patterns of recurrence (1). Thus, after having settled the original problem concerning the Poisson–Dirichlet coalescent, our method was subsequently extended to cover many other linear recurrences (1).

In this paper, unless stated the contrary, we tacitly suppose that $b_n \geq 0$ and, hence, $a_n \geq 0$. However, a perusal of the proofs given below reveals that we could have assumed that b_n is only non-negative or non-positive for large enough n. Under this last assumption, the formulations of results would get cumbersome which has forced us to keep less generality but more transparency.

Our method can be summarized in the following

ALGORITHM

(1) Using, for example, the method described in [28], obtain the recurrence with weights reduced to probabilities. As a result, we obtain a recurrence of the form

$$A_1 = 0, \quad A_n = B_n + \sum_{k=1}^{n-1} p_{nk} A_k,$$

where $\sum_{k=1}^{n-1} p_{nk} = 1$ for all $n \geq 2$ and $B_n \geq 0$. Let I_n be a random variable with the distribution $\mathbb{P}\{I_n = k\} = p_{nk}, n \geq 2, k < n$.

(2) Prove the divergence of A_n using, for example, Proposition 5.1 or other methods.

(3) Find a continuous, strictly increasing, and unbounded function $g(x)$ defined on $\mathbb{R}^+$ and such that $g(n) = \mathbb{E}I_n + o(\mathbb{E}I_n)$. Pick x_0 as defined in (3). Find a continuous function $h(x)$ defined on $\mathbb{R}^+$ and such that $h(n) = B_n$.

(4) Find an iterative function g^* generated by the quadruple (h, g, x_0, k), where k is any function continuous on $[0, x_0]$ (see Definition 2.1).

(5) Using, for example, Theorem 5.2, find, if possible, an elementary function f_1 such that $\lim_{x\to\infty} \frac{f_1(x)}{g^*(x)} = 1$, and set $f := f_1$. Otherwise, select k such that g^* is twice differentiable, and set $f := g^*$ (see Theorem 2.1).

(6) If $f(\mathbb{E}I_n) - f(g(n)) = o(h(n))$, then go to the next step, otherwise go to step 3) and choose an asymptotically smaller term $o(\mathbb{E}I_n)$.

[1]The result for $\mathbb{E}X_n$ was conjectured by M. Möhle.

(7) If $\mathbb{E}f(I_n) - f(\mathbb{E}I_n) = o(h(n))$, then $A_n \sim f(n)$. If $\mathbb{E}f(I_n) - f(\mathbb{E}I_n) \sim ch(n)$, then $A_n \sim (1-c)^{-1} f(n)$ (see Theorem 3.2).

We mention, in passing, that iterative functions have already been used in the context of the divide-and-conquer paradigm [19]. The cited paper is concerned with the stochastic processes $\{T(x), x \in \mathbb{R}^+\}$, whose marginal distributions are given by the equality

$$T(x) \stackrel{d}{=} a(x) + T'(t(x)), \quad x \in \mathbb{R}^+,$$

where $a(\cdot)$ is a non-negative (deterministic) function, and $t(\cdot)$ is a random variable taking values in $[0, \cdot]$ which is independent of $\{T'(x), x \in \mathbb{R}^+\}$, an independent copy of $\{T(x), x \in \mathbb{R}^+\}$.

The structure of the paper is as follows. Section 2 introduces iterative functions and investigates their properties. Section 3 carefully describes the algorithm of our new method. Theorem 3.1 and Theorem 3.2 which are the main results of the section prove the validity of the algorithm. Section 4 is devoted to applications and also discusses "ins and outs" of the method. The paper closes with Appendix which collects proofs of some technical results concerning the iterative functions and properties of recurrences (2).

Throughout the paper, the notation $r(\cdot) \sim s(\cdot)$ means that $r(\cdot)/s(\cdot) \to 1$, as the argument tends to infinity, $C^{(m)}(B)$ denotes the space of functions which are m-times differentiable on the set B. If $B = [a, \infty)$, then the derivatives at point a are assumed to be the right derivatives. Also we use the notation

$$r^{\circ(0)}(x) \stackrel{def}{=} x, \quad r^{\circ(k)}(x) \stackrel{def}{=} r(r^{\circ(k-1)}(x)), \quad k \in \mathbb{N}.$$

Finally, we recall the standard notation $\lfloor x \rfloor = \sup\{k \in \mathbb{Z} : k \le x\}$ and $\lceil x \rceil = \inf\{k \in \mathbb{Z} : k \ge x\}$ for the floor and ceiling functions, respectively.

2. Iterative functions

In this section, iterative functions are defined, and some basic properties of these functions are given. We start with a formal definition.

Definition 2.1. Suppose that the function $g : \mathbb{R}^+ \to \mathbb{R}^+$ is increasing, unbounded, and continuous and satisfies the following condition: for some $x_0 > 0$ and every $x_1 > x_0$, there exists $\varepsilon_{x_1} > 0$ such that

$$x - g(x) > \varepsilon_{x_1} \quad \text{for all} \quad x \in (x_0, x_1). \tag{3}$$

Assume that the functions $h : \mathbb{R}^+ \to \mathbb{R}^+$ and $k : [0, x_0] \to \mathbb{R}$ are continuous and define the function $g^* : \mathbb{R}^+ \to \mathbb{R}$ by the equality

$$g^*(x) = \sum_{i=1}^{m_0(x)} h(g^{\circ(i-1)}(x)) + k(g^{\circ(m_0(x))}(x)), \tag{4}$$

where

$$m_0(x) := \inf\{k \ge 0 : g^{\circ(k)}(x) \le x_0\}.$$

We call g^* the *iterative function generated by the quadruple* (h, g, x_0, k) and denote it by $g^* = \mathrm{Iter}(h, g, x_0, k)$.

Note that the technical condition (3) is sufficient for $m_0(x)$ to be finite for every $x \in \mathbb{R}^+$. This follows from the estimate $m_0(x) \le \lfloor \frac{x - x_0}{\varepsilon_x} \rfloor + 1$, which is implied by the inequality

$$x - k\varepsilon_x \ge g^{\circ(k)}(x), \quad x > x_0, \quad k = 0, \dots, m_0(x).$$

which can be obtained, in turn, by induction.

Remark 2.1. It follows from the definition that g^* satisfies the functional equation

$$g^*(x) = h(x) + g^*(g(x)), \; x > x_0, \tag{5}$$

with the initial condition

$$g^*(x) = k(x), \qquad x \le x_0.$$

Below are some examples of iterative functions

Example 1. Let $h(x) \equiv 1$, $g(x) = \alpha x$, $\alpha \in [0,1)$, $x_0 = 1$, $k(x) \equiv 0$. Then $g^*(x) = 1 + g^*(\alpha x)$, $x > 1$, or

$$g^*(x) = \lceil \log_{\frac{1}{\alpha}} x \rceil, \; x > 1.$$

Example 2. Let $h(x) \equiv 1$, $g(x) = \log x$, $x_0 = 1$, $k(x) \equiv 0$. Then $g^*(x) = 1 + g^*(\log x)$, $x > 1$, or

$$g^*(x) = \log^* x,$$

the *log-star function* which is arguably the best known non-trivial iterative function. It is clear that $\mathrm{Iter}(1, g, x_0, 0) = m_0(x)$. In particular, this equality holds for the log-star function.

If $h(x_0) \ne 0$, then the iterative functions $\mathrm{Iter}(h, g, x_0, 0)$ are piecewise continuous. We prefer, however, to work with smooth iterative functions, which was the main reason for introducing the functions k in Definition 2.1. It turns out that $\mathrm{Iter}(h, g, x_0, 0)$ and $\mathrm{Iter}(h, g, x_0, k)$ have the same asymptotics, and an appropriate choice of k makes $\mathrm{Iter}(h, g, x_0, k)$ smooth enough. Below, we formalize this statement and also describe how the mentioned smoothness can be obtained by the choice of k.

Introduce the equivalence relation $\approx$ on the set of iterative functions by the rule

$$g_1^* \approx g_2^* \iff g_1^* = \mathrm{Iter}(h, g, x_0, k_1), \;\; g_2^* = \mathrm{Iter}(h, g, x_0, k_2).$$

This relation induces partitioning the set of iterative functions into the classes of equivalence.

Definition 2.2. The equivalence class

$$\mathcal{F} := \{F = \mathrm{Iter}(h, g, x_0, k), k \in C[0, x_0]\}$$

is called the *iterative function generated by the triple* (h, g, x_0). When it does not lead to ambiguity, we call an *iterative function generated by the triple* (h, g, x_0) an arbitrary element of this class.

Since $|g_1^*(x) - g_2^*(x)|$ is bounded on $\mathbb{R}^+$, for any $g_1^*, g_2^* \in \mathcal{F}$, all iterative functions in the same equivalence class are asymptotically equivalent (provided they diverge).

Definition 2.3. An *m-time differentiable modification* of the iterative function g^* is an arbitrary iterative function $\hat{g}^*$ such that $\hat{g}^* \approx g^*$ and $\hat{g}^* \in C^{(m)}[x_0, +\infty)$.

Our first result which is a direct consequence of Lemma 5.1 and Lemma 5.2 given in the Appendix shows that, provided g and h are smooth enough, one can find a function k such that the function $\mathrm{Iter}(h, g, x_0, k)$ is smooth. For a collection of functions $f_1, \ldots, f_n$, let $W(f_1, \ldots, f_n)$ denote its Wronskian.

Theorem 2.1. *Assume that* $g, h \in C^{(m)}[x_0, +\infty)$ *and that*

$$W\Big(x^i - g^i(x), i = 0, \ldots, m+1\Big)(x_0) \ne 0.$$

Then there exists a function k *of the form*

$$k(x) = \sum_{i=1}^{m+1} \alpha_i x^i,$$

such that the iterative function generated by the quadruple (h, g, x_0, k) *is* m*-time differentiable on* $[x_0, +\infty)$.

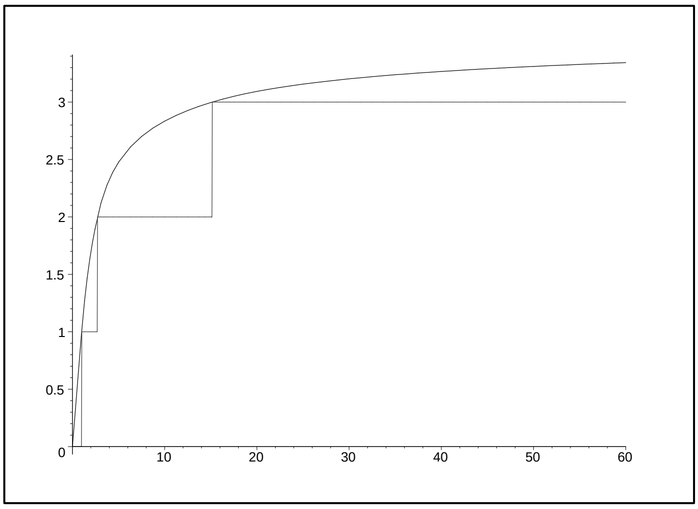
Remark 2.2. The vector of coefficients $(\alpha_1, \alpha_2, \dots, \alpha_{m+1})$ is a solution to the system of linear equations (see Lemma 5.2) and can be calculated explicitly.

An example of a smoothed iterative function is given as follows.

Example 2.1. Recall that the log-star function is an iterative function generated by the quadruple $(1, \log x, 1, 0)$. A twice differentiable modification F of the log-star function can be constructed in the following way. According to Lemma 5.2, the corresponding function k takes the form $k(x) = -\frac{2}{13}x^3 + \frac{3}{13}x^2 + \frac{12}{13}x$. Therefore,

$$F(x) = \begin{cases} 1 + F(\log x), & x > 1, \\ -\frac{2}{13}x^3 + \frac{3}{13}x^2 + \frac{12}{13}x, x \in [0, 1]. \end{cases}$$

Below are depicted the graphs of the functions $\log^* x$ and $F(x)$ for $x > 0$.



3. Asymptotic behavior of (2)

While investigating recurrence (2), without loss of generality, we can assume that, for every $n \geq 2$,

$$\sum_{k=1}^{n-1} c_{nk} = 1 \text{ and } c_{nk} \geq 0, \ k = 1, \dots, n-1 \tag{6}$$

(see, e.g., p. 9 in [28]). In what follows, recurrences (2) with $b_n \geq 0$ which satisfy (6) are referred to as *recurrences with weights reduced to probabilities.* If (6) holds, we denote, by I_n, a random variable with the distribution

$$\mathbb{P}\{I_n = k\} = c_{nk}, \quad k = 1, \dots, n-1.$$

Theorem 3.1. *Assume that the sequence $\{a_n, n \in \mathbb{N}\}$ satisfy recurrence* (2) *with weights reduced to probabilities. Let $g : \mathbb{R}^+ \to \mathbb{R}^+$ be a continuous, increasing, and unbounded function such that*

$$g(n) = \mathbb{E}I_n + o(\mathbb{E}I_n), \quad n \to \infty,$$

and let $h : \mathbb{R}^+ \to \mathbb{R}^+$ be a continuous function such that

$$h(n) = b_n, \quad n \geq 2.$$

If

- $\lim_{n\to\infty} a_n = +\infty$,
- $g^*(\mathbb{E}I_n) - g^*(g(n)) = o(h(n))$, $n \to \infty$,

where g^* is an iterative function generated by the triple (h, g, x_0), then the following implications are true

$$\mathbb{E}g^*(I_n) - g^*(\mathbb{E}I_n) = o(h(n)), \quad n \to \infty \Longrightarrow \\ a_n \sim g^*(n), \quad n \to \infty, \tag{7}$$

$$\mathbb{E}g^*(I_n) - g^*(\mathbb{E}I_n) \sim dh(n), \quad n \to \infty, \quad \text{for some } d < 1 \Longrightarrow \\ a_n \sim (1-d)^{-1}g^*(n), \quad n \to \infty. \tag{8}$$

Proof. Set $a'_n := a_n - g^*(n)$, $n \in \mathbb{N}$. The sequence $\{a'_n, n \in \mathbb{N}\}$ satisfies the recurrence

$$a'_1 = -g^*(1), \quad a'_n = b_n - g^*(n) + \mathbb{E}g^*(I_n) + \sum_{k=1}^{n-1} c_{nk}a'_k, \quad n \geq 2. \tag{9}$$

If $\mathbb{E}g^*(I_n) - g^*(\mathbb{E}I_n) = o(h(n))$ and $g^*(\mathbb{E}I_n) - g^*(g(n)) = o(h(n))$, then the inhomogeneous term of (9) is $o(h(n))$. Therefore, applying part (II) of Theorem 5.1 yields $a'_n = o(a_n)$ which implies $a_n \sim g^*(n)$.

If $\mathbb{E}g^*(I_n) - g^*(\mathbb{E}I_n) \sim dh(n)$ for some $d \in (0, 1)$ and $g^*(\mathbb{E}I_n) - g^*(g(n)) = o(h(n))$, then the inhomogeneous term of (9) is asymptotically equal to $dh(n)$. Therefore, applying part (I) of Theorem 5.1 yields $a'_n \sim da_n$ which implies $a_n \sim (1-d)^{-1}g^*(n)$.

Finally, if $\mathbb{E}g^*(I_n) - g^*(\mathbb{E}I_n) \sim dh(n)$ for some $d < 0$, we can apply part (II) of Theorem 5.1 to the sequences $\{g^*(n) - a_n\}$ and $\{a_n\}$ to conclude that $g^*(n) - a_n \sim -da_n$. The latter is equivalent to $a_n \sim (1-d)^{-1}g^*(n)$. The proof is complete. □

Theorem 3.2. *Assume that the sequence $\{a_n, n \in \mathbb{N}\}$ satisfy recurrence* (2) *with weights reduced to probabilities. Let $g : \mathbb{R}^+ \to \mathbb{R}^+$ be a twice differentiable, increasing, and unbounded function such that*

$$g(n) = \mathbb{E}I_n + o(\mathbb{E}I_n), \quad n \to \infty,$$

and $h : \mathbb{R}^+ \to \mathbb{R}^+$ be a twice differentiable function such that

$$h(n) = b_n, \quad n \geq 2.$$

If the conditions

(C1) $\lim_{n\to\infty} a_n = +\infty$
(C2) *There exists a continuous function k such that the iterative function F generated by the quadruple (h, g, x_0, k) is twice differentiable*
(C3) $F(\mathbb{E}I_n) - F(g(n)) = o(h(n))$, $n \to \infty$
(C4) *There exists $M > 0$ such that for all $n \in \mathbb{N}$*

$$\operatorname{Var} I_n \leq M\mathbb{E}I_n$$

(C5) $\lim_{n\to\infty} \sup_{x \geq \mathbb{E}I_n/2} |F''(x)| \dfrac{\operatorname{Var} I_n}{h(n)} = 0$
(C6) $\lim_{n\to\infty} \dfrac{\sup_{1\leq x\leq n} |F(x)|}{h(n)\operatorname{Var} I_n} = 0$
(C7) $\lim_{n\to\infty} \dfrac{F'(\mathbb{E}I_n)}{h(n)} = 0$

hold, then

$$a_n \sim F(n), \; n \to \infty.$$

Proof. Since conditions (C1) and (C3) hold, according to implication (7) in Theorem 3.1, it is enough to show that

$$\alpha_n := \mathbb{E}F(I_n) - F(\mathbb{E}I_n) = o(h(n)).$$

With $\kappa := \frac{1}{2M}$ and $A_n := \{|I_n - \mathbb{E}I_n| > \kappa \mathrm{Var}\, I_n\}$, we have

$$|\alpha_n| \le |\mathbb{E}(F(I_n) - F(\mathbb{E}I_n))1_{A_n}| + |\mathbb{E}(F(I_n) - F(\mathbb{E}I_n))1_{A_n^c}| =: \beta_n + \gamma_n.$$

The application of Chebyshev's inequality yields

$$\beta_n \le 2 \sup_{1\le x\le n} |F(x)|\mathbb{P}(A_n) \le \frac{2\mathrm{Var}\, I_n \sup_{1\le x\le n}|F(x)|}{(\kappa \mathrm{Var}\, I_n)^2},$$

which is $o(h(n))$ by condition (C6).

Using the Taylor expansion around $\mathbb{E}I_n$ leads to

$$\begin{aligned}\gamma_n &= \Big|\mathbb{E}\Big(F'(\mathbb{E}I_n)(I_n - \mathbb{E}I_n) + \frac{1}{2}F''(\theta_n)(I_n - \mathbb{E}I_n)^2\Big)1_{A_n^c}\Big| \\ &\le \Big|F'(\mathbb{E}I_n)\mathbb{E}(I_n - \mathbb{E}I_n)1_{A_n}\Big| + \frac{1}{2}\Big|\mathbb{E}F''(\theta_n)(I_n - \mathbb{E}I_n)^2 1_{A_n^c}\Big| = \gamma_{1,n} + \gamma_{2,n},\end{aligned}$$

where $\theta_n \in [\mathbb{E}I_n - \kappa \mathrm{Var}\, I_n, \mathbb{E}I_n + \kappa \mathrm{Var}\, I_n]$. Consequently, by Cauchy–Schwarz and Chebyshev's inequalities, we obtain

$$\begin{aligned}\gamma_{1,n} &= |F'(\mathbb{E}I_n)\mathbb{E}(I_n - \mathbb{E}I_n)1_{A_n}| \le |F'(\mathbb{E}I_n)|\sqrt{\mathbb{E}(I_n - \mathbb{E}I_n)^2}\sqrt{\mathbb{P}(A_n)} \\ &\le |F'(\mathbb{E}I_n)|\sqrt{\mathrm{Var}\, I_n}\sqrt{\frac{\mathrm{Var}\, I_n}{(\kappa \mathrm{Var}\, I_n)^2}} = \frac{1}{\kappa}|F'(\mathbb{E}I_n)|,\end{aligned}$$

which is $o(h(n))$ by condition (C7).

Finally, the appeal to condition (C4) allows us to conclude that

$$S_4 \le \frac{1}{2} \sup_{x\ge \mathbb{E}I_n/2} |F''(x)|\mathrm{Var}\, I_n,$$

which is $o(h(n))$ by condition (C5). The proof is complete. □

Theorem 3.1 and Theorem 3.2 justify the algorithm given in Introduction.

4. Applications

4.1. **Exchangeable coalescents.**

4.1.1. *Number of collisions in beta$(a,1)$-coalescents.* Let X_n be the number of collisions in beta$(a,1)$-coalescent, $a > 0$, restricted to the set $\{1,\dots,n\}$. Many results concerning the asymptotics of $\mathbb{E}X_n^k$, $k \in \mathbb{N}$, are known [3, 6, 7, 12, 16, 22, 25], but we partially derive them again just in order to show how our method works.

It is known (see, e.g., [17, Section 7]) that the sequence $\{X_n, n \in \mathbb{N}\}$ satisfies the distributional equality

$$X_1 = 0, \quad X_n \overset{d}{=} 1 + X_{I_n}, \quad n \ge 2,$$

where I_n is a random variable with the distribution

$$\mathbb{P}\{I_n = n-k\} = \frac{\frac{(2-a)\Gamma(a+k-1)}{\Gamma(a)\Gamma(k+2)}}{1 - \frac{\Gamma(a+n-1)}{\Gamma(a)\Gamma(n+1)}}, \quad k = 1,\dots,n-1, \quad n \ge 2,$$

if $a \ne 2$, and

$$\mathbb{P}\{I_n = n-k\} = \frac{1}{(h_n - 1)(k+1)}, \quad k = 1,\dots,n-1, \quad n \ge 2,$$

where $h_n = \sum_{k=1}^n k^{-1}$, if $a = 2$.

By Proposition 5.1, it follows that $\lim_{n\to\infty} \mathbb{E}X_n = +\infty$. It is also clear that no reduction of weights to probabilities in the recurrence is needed.

Case $0 < a < 1$ [3, 12, 17]. Since

$$\mathbb{E}I_n = n - (1-a)^{-1} + o(1),$$

we can choose

$$g(x) = x - \frac{1}{1-a} \quad \text{and} \quad h(x) = 1.$$

Then the functional equation (5) has an elementary solution $g^*(x) = (1-a)x$. By Theorem 3.1, $\mathbb{E}X_n \sim g^*(n) \sim (1-a)n$.

Case $a = 1$ (Bolthausen–Sznitman coalescent) [6, 17, 22, 25]. Since

$$\mathbb{E}I_n = n - \log n + O(1),$$

we can choose

$$g(x) = x - \log x \quad \text{and} \quad h(x) = 1.$$

From the relation $\frac{x}{\log x} = 1 + o(1) + \frac{x - \log x}{\log(x - \log x)}$ and Theorem 5.2, it follows that $\mathrm{Iter}(h, g, 2)(x) \sim \frac{x}{\log x}$. The application of Theorem 3.2[2] gives $\mathbb{E}X_n \sim \frac{n}{\log n}$.

Case $a = 2$ [16]. Since

$$\mathbb{E}I_n = n - \frac{n}{\log n} + O\left(\frac{n}{\log^2 n}\right),$$

we can choose

$$g(x) = \left(x - \frac{x}{\log x}\right) 1_{(e,\infty)}(x) \quad \text{and} \quad h(x) = 1.$$

From the relation $\frac{1}{2}\log^2 x = 1 + o(1) + \frac{1}{2}\log^2\left(x - \frac{x}{\log x}\right)$ and Theorem 5.2, we conclude that $\mathrm{Iter}(h, g, 2)(x) \sim \frac{1}{2}\log^2 x$. Direct calculations show that

$$\log^2 \mathbb{E}I_n - \log^2 g(n) = O\left(\frac{1}{\log n}\right)$$

and

$$\lim_{n\to\infty} \frac{1}{2}\left(\mathbb{E}\log^2 I_n - \log^2 \mathbb{E}I_n\right) = 1 - \frac{\pi^2}{6},$$

which, in view of implication (8), yields $\mathbb{E}X_n \sim \frac{3}{\pi^2}\log^2 n$.

4.1.2. *Functionals acting on the Poisson–Dirichlet coalescent.* Unlike the beta coalescents, the asymptotics of the moments of the number of collisions X_n in the Poisson–Dirichlet coalescent do not seem to have been known so far (we refer to [23, 29] for the extensive information about this particular coalescent with simultaneous multiple collisions). We recall that this fact has served as an initial motivation for developing the method reported in this article.

It can be checked that the sequence $\{X_n, n \in \mathbb{N}\}$ satisfies the distributional equality

$$X_1 = 0, \quad X_n \overset{d}{=} 1 + X_{I_n}, \quad n \geq 2,$$

where

$$\mathbb{P}\{I_n = k\} = \frac{\theta^k}{[\theta]_n - \theta^n} s(n,k), \quad k = 1, \dots, n-1, \quad n \geq 2,$$

$s(n,k)$ is the unsigned Stirling number of the first kind, and $[\theta]_n = \theta(\theta+1)\dots(\theta+n-1)$. This implies that

$$\mathbb{E}I_n = \theta \log n + O(1), \quad \mathrm{Var}\, I_n = \theta \log n + O(1).$$

[2]The only thing which may require a verification is condition $C3$. In the present situation, $\mathbb{E}I_n - g(n) = O(1)$, and the derivative of $F(x) = x/\log x$ tends to zero, as $x \to \infty$. Therefore, condition $C3$ follows by the application of the mean value theorem.

Set

$$g(x) := \theta \log x, \quad h(x) = 1 \ \text{ and } \ g^*(x) = \mathrm{Iter}(h, g, x_0)$$

for some $x_0 > \exp(2\theta \vee 1)$. Notice that g^* is a generalized log-star function which can be defined via the functional equation

$$g^*(x) = 1 + g^*(\theta \log x), \ x > x_0.$$

Let F be a twice differentiable modification of g^* of the form

$$F(x) = \begin{cases} 1 + F(\theta \log x), & x > x_0, \\ \alpha_1 x^3 + \alpha_2 x^2 + \alpha_3 x, x \in [0, x_0], \end{cases}$$

for some constants $\alpha_1, \alpha_2, \alpha_3$. It follows that, for every fixed $j \in \mathbb{N}$,

$$F'(x) = o\bigg(\frac{1}{x \log x \cdots \log^{\circ(j)}(x)}\bigg) \ \text{ and } \ F''(x) = o\bigg(\frac{1}{x^2 (\log x)^2 \cdots (\log^{\circ(j)}(x))^2}\bigg).$$

The application of Theorem 3.2 yields $\mathbb{E}X_n \sim g^*(n) \sim F(n)$. Analogously, we obtain

$$\mathbb{E}X_n^k \sim (g^*(n))^k, \quad k \in \mathbb{N}.$$

Other important functionals acting on the Poisson–Dirichlet coalescent are the absorption time T_n (in the biological context, T_n is the time back to the most recent common ancestor of a sample of size n) and the total branch length L_n of the coalescent tree.

The corresponding distributional recurrences are

$$T_1 = 0, \ \ T_n \overset{d}{=} \tau_n + T_{I_n}, \ \ n \geq 2,$$
$$L_1 = 0, \ \ L_n \overset{d}{=} n\tau_n + L_{I_n}, \ \ n \geq 2,$$

where τ_n is a random variable with the exponential law with the parameter $g_n = 1 - \frac{\theta^n}{[\theta]_n}$ which is independent of everything else.

Using induction on k, the fact that $\lim_{n\to\infty} g_n = 1$, and Theorem 5.1, we conclude that

$$\mathbb{E}T_n^k \sim (g^*(n))^k, \quad k \in \mathbb{N}.$$

The application of Chebyshev's inequality immediately leads to the following weak laws of large numbers.

Theorem 4.1. *As $n \to \infty$,*

$$\frac{X_n}{g^*(n)} \overset{P}{\to} 1 \ \text{ and } \ \frac{T_n}{g^*(n)} \overset{P}{\to} 1.$$

As far as L_n is concerned, we can prove that

$$\mathbb{E}L_n^k \sim k! n^k, \quad k \in \mathbb{N}.$$

By the method of moments, this immediately gives the following weak convergence result.

Proposition 4.1. *As $n \to \infty$,*

$$\frac{L_n}{n} \overset{d}{\to} L,$$

where L is a random variable with the standard exponential law.

In a recent preprint [23], the same result was obtained by a different method. We thus omit further details.

4.2. **Examples from the analysis of algorithms.** We will give new proofs of the results from [24], [21], and [27], respectively, by using our method.

4.2.1. *The Quickselect algorithm.* Let X_n be the number of comparisons that the Quickselect algorithm needs to find $\min(x_1, \dots, x_n)$ of a sample $x_1, \dots, x_n$. Then

$$X_1 = 0, \quad X_n \overset{d}{=} n - 1 + X_{I_n}, \quad n \geq 2,$$

where $I_n = J_n \vee 1$, and J_n is uniformly distributed on $\{0, \dots, n-1\}$. Since

$$\mathbb{E} I_n = \frac{n-1}{2} + \frac{1}{n},$$

we can choose

$$g(x) = \frac{x+1}{2} \quad \text{and} \quad h(x) = x - 1.$$

Then the functional equation (5) has elementary solutions $g^*(x) = 2x + c$, $c \in \mathbb{R}$. By Theorem 3.1, $\mathbb{E}X_n \sim g^*(n) \sim 2n$.

4.2.2. *The depth of a random node in a random binary search tree.* The corresponding recurrence is

$$X_0 = -1, \; X_1 = 0, \quad X_n \overset{d}{=} 1 + X_{I_n}, \quad n \geq 2,$$

where $\mathbb{P}\{I_n = k\} = 2k/n^2$ for $k \in \{1, \dots, n-1\}$ and $\mathbb{P}\{I_n = 0\} = 1/n$. Since

$$\mathbb{E} I_n = \frac{(n-1)(2n-1)}{3n},$$

we can choose

$$g(x) = 2x/3 \quad \text{and} \quad h(x) = 1.$$

According to Example 1, the corresponding iterative function is

$$g^*(x) = \lceil \log_{\frac{3}{2}} x \rceil, \quad x > 1.$$

Since $\lim_{n\to\infty}(\mathbb{E}\log^+ I_n - \log n) = -1/2$, it follows that $\lim_{n\to\infty}(\mathbb{E}f(I_n) - f(\mathbb{E}I_n)) = 1 - \frac{1}{2\log(3/2)}$, where

$$f(x) = \frac{\log^+ x}{\log(3/2)}.$$

Since $f(x) \sim g^*(x)$ then, according to the algorithm, $\mathbb{E}X_n \sim 2\log(3/2) f(n) \sim 2\log n$.

4.2.3. *The Quicksort algorithm.* Let X_n denote the random number of comparisons needed to sort a list of length n by the Quicksort. Then $X_0 = X_1 = 0$, and

$$X_n \overset{d}{=} n - 1 + X_{I_n - 1} + X'_{n - I_n}, \quad n \geq 2,$$

where $\{X'_n, n \in \mathbb{N}_0\}$ is an independent copy of $\{X_n, n \in \mathbb{N}_0\}$ which is independent of I_n having the uniform distribution on $\{1, \dots, n\}$. Set $a_n := \mathbb{E}X_n$. Then $a_0 = a_1 = 0$ and

$$a_n = n - 1 + \sum_{k=0}^{n-1} \frac{2}{n} a_k, \quad n \geq 2.$$

The reduction of weights to probabilities can be made by the substitution $a'_n := a_n/(n+1)$ which yields

$$a'_n = \frac{n-1}{n+1} + \sum_{k=0}^{n-1} \frac{2(k+1)}{n(n+1)} a'_k, \quad n \geq 2.$$

Using the same arguments as in the previous example, we obtain $a'_n \sim 2\log n$. Therefore, $\mathbb{E}X_n \sim 2n\log n$, which is the well-known asymptotics for the Quicksort. [3]

[3]The first result concerning the complexity of the (non-randomized) Quicksort algorithm with $O(n \log n)$ asymptotics goes back to the pioneer work by Hoar [14]. For a complete analysis of the Quicksort and its different modifications, we refer to survey [30].

4.3. **Limitations of the method.**

(a) An indispensable requirement of our method to work is the divergence of a_n, the solution to (2). In particular, our method cannot detect the convergence of a_n to a constant.
(b) It may be difficult to guess which elementary function has the same asymptotics as a given iterative function.
(c) If condition (8) holds for some $c \neq 0$, it may be hard to calculate the constant c explicitly. Therefore, it seems that a natural assumption for the method to work is (7) rather than (8). Condition (7) holds if the solution is nearly linear and the variance of index I_n grows not too fast (precise statements are made in Theorem 3.2). For instance, the mean number of collisions in the Bolthausen–Sznitman and Poisson–Dirichlet coalescents exhibit the asymptotic behavior of this type.

5. Appendix.

5.1. **Some properties of iterative functions.** For the given strictly increasing continuous function g, there exists the unique inverse function g^{-1} which defines the sequence $\{A_n, n \in \mathbb{N}_0\}$ as follows:

$$A_0 = 0, \quad A_i := (g^{-1})^{\circ(i-1)}(x_0), \quad i \in \mathbb{N}. \tag{10}$$

Lemma 5.1. *Assume that* $g, h, k \in C^{(m)}[x_0, +\infty)$, *and* $F = \mathrm{Iter}(h, g, x_0, k)$ *is* m*-time differentiable at* x_0. *Then* F *is* m*-time differentiable on* $[x_0, +\infty)$.

Proof. We only treat the case $m = 1$, as, for $m = 2, 3, \ldots$, the proof is the same. Since F is a sum of compositions of $C^{(1)}[x_0, +\infty)$ functions, it is differentiable on $[x_0, +\infty)\backslash\{A_i, i \in \mathbb{N}\}$. Therefore, we only have to check the continuity and the differentiability at points $\{A_i,\ i \in \mathbb{N}\}$.

First step. Proof of the continuity. By assumption, F is continuous at $A_1 = x_0$, i.e.,

$$k(x_0) = h(x_0) + k(g(x_0)). \tag{11}$$

For fixed $k \geq 2$, (4) yields

$$F(A_k - 0) = \sum_{i=1}^{k-1} h(g^{\circ(i-1)}(A_k - 0)) + k(g^{\circ(k-1)}(A_k - 0)) \tag{12}$$

and

$$F(A_k + 0) = \sum_{i=1}^{k} h(g^{\circ(i-1)}(A_k + 0)) + k(g^{\circ(k)}(A_k + 0)). \tag{13}$$

Use now (11) and the continuity of h and g to obtain

$$\begin{aligned} &F(A_k + 0) - F(A_k - 0) \\ &\quad = h(g^{\circ(k-1)}(A_k)) + k(g^{\circ(k)}(A_k)) - k(g^{\circ(k-1)}(A_k)) \\ &\quad = h(x_0) + k(g(x_0)) - k(x_0) \overset{(11)}{=} 0. \end{aligned}$$

Second step. Proof of the differentiability. The differentiability of F at x_0 implies that

$$k'(x_0) = h'(x_0) + k'(g(x_0))g'(x_0). \tag{14}$$

For $k \geq 2$, using (12) and (13) yields

$$F'_{-}(A_k) = \lim_{x \to A_k - 0} \frac{d}{dx}\biggl(\sum_{i=1}^{k-1} h(g^{\circ(i-1)}(x)) + k(g^{\circ(k-1)}(x))\biggr),$$

$$F'_{+}(A_k) = \lim_{x \to A_k + 0} \frac{d}{dx}\biggl(\sum_{i=1}^{k} h(g^{\circ(i-1)}(x)) + k(g^{\circ(k)}(x))\biggr).$$

Consequently,

$$\begin{aligned} &F'_{+}(A_k) - F'_{-}(A_k) \\ &= \lim_{x \to A_k + 0} \frac{d}{dx} h(g^{\circ(k-1)}(x)) + k(g^{\circ(k)}(x)) - \lim_{x \to A_k - 0} \frac{d}{dx} k(g^{\circ(k-1)}(x)). \end{aligned}$$

Set $u(x) := g^{\circ(k-1)}(x)$. Then $u(A_k + 0) = u(A_k - 0) = u(A_k) = x_0$ and

$$\begin{aligned} &F'_{+}(A_k) - F'_{-}(A_k) \\ &= \lim_{x \to A_k + 0} \frac{d}{dx} h(u(x)) + k(g(u(x))) - \lim_{x \to A_k - 0} \frac{d}{dx} k(u(x)) \\ &= \lim_{x \to A_k + 0} (h'(u(x)) + k'(g(u(x)))g'(u(x)))u'(x) - \lim_{x \to A_k - 0} k'(u(x)))u'(x) \\ &= (h'(x_0) + k'(g(x_0))g'(x_0) - k'(x_0))u'(x_0) = 0, \end{aligned}$$

by (14). The proof is complete. □

From this lemma, it follows that the function F is m-time differentiable, provided it satisfies the conditions

$$\begin{aligned} &k(x_0) = h(x_0) + k(g(x_0)), \\ &k'(x_0) = h'(x_0) + k'(g(x_0))g'(x_0), \\ &\dots\dots\dots\dots\dots\dots\dots\dots \\ &k^{(m)}(x_0) = h^{(m)}(x_0) + (k(g(x_0)))^{(m)}. \end{aligned} \tag{15}$$

The following lemma proves the existence of such a function $k(x)$.

Lemma 5.2. *Assume that* $W\Bigl(x - g(x), \dots, x^{m+1} - g^{m+1}(x)\Bigr)\Big|_{x=x_0} \neq 0$. *Then there exists a function* $k(x) = \sum_{i=1}^{m+1} \alpha_i x^i$ *which satisfies (15).*

Proof. Plugging the representation $k(x) = \sum_{i=1}^{m+1} \alpha_i x^i$ into (15) gives the system of linear equations

$$\begin{aligned} &\Bigl(\alpha_1(x_0 - g(x_0)) + \dots + \alpha_{m+1}(x_0^{m+1} - g^{m+1}(x_0))\Bigr) = h(x_0), \\ &\Bigl(\alpha_1 \tfrac{d}{dx}(x - g(x)) + \dots + \alpha_{m+1}\tfrac{d}{dx}(x^{1+m} - g^{m+1}(x))\Bigr)\Big|_{x=x_0} = h'(x_0), \\ &\dots\dots\dots\dots\dots\dots\dots\dots \\ &\Bigl(\alpha_1 \tfrac{d^m}{dx^m}(x - g(x)) + \dots + \alpha_{m+1}\tfrac{d^m}{dx^m}(x^{m+1} - g^{m+1}(x))\Bigr)\Big|_{x=x_0} = h^{(m)}(x_0). \end{aligned}$$

The determinant of this system is $W(x_0)$ which is not equal to zero by assumption. Therefore, the system has a unique solution which implies that the function k is well defined and satisfies conditions (15). □

5.2. **Inhomogeneous terms of recursion (2) and iterative functions.**

Theorem 5.1. *Suppose that* $\{a_n, n \in \mathbb{N}\}$ *and* $\{a'_n, n \in \mathbb{N}\}$ *satisfy the recurrences*

$$a_n = b_n + \sum_{k=1}^{n-1} p_{nk} a_k, n \geq N \tag{16}$$

and

$$a'_n = b'_n + \sum_{k=1}^{n-1} p_{nk} a'_k, n \geq N, \tag{17}$$

respectively. Suppose that $b_n \geq 0$ *for* $n \geq N$ *and* $\lim_{n\to\infty} a_n = +\infty$*. Then*

I. $b'_n \sim b_n$, $n \to \infty$ *implies* $a'_n \sim a_n$, $n \to \infty$, *and*
II. $b'_n = o(b_n)$, $n \to \infty$ *implies* $a'_n = o(a_n)$, $n \to \infty$.

Proof of (I). We exploit the idea of the proof of [9, Proposition 3]. Suppose there exists $\varepsilon_0 > 0$ such that $a_n > (1+\varepsilon_0)a'_n$ for infinitely many n. Since $\lim_{n\to\infty} a_n = +\infty$, we can pick $\varepsilon \in (0, \varepsilon_0]$ such that, for any $c > 0$, the inequality $a_n > (1+\varepsilon)a'_n + c$ holds for infinitely many n. Let n_c be such minimal n. Since $\lim_{c\to\infty} n_c = +\infty$, we can assume, without loss of generality, that $n_c > N$. For $n \leq n_c - 1$, we have $a_n < (1+\varepsilon)a'_n + c$, which implies

$$(1+\varepsilon)a'_{n_c} + c < a_{n_c} = b_{n_c} + \sum_{k=1}^{n_c-1} p_{n_c k} a_k < b_{n_c} + c + (1+\varepsilon)\sum_{k=1}^{n_c-1} p_{n_c k} a'_k.$$

Simplifying the last expression gives $1 + \varepsilon < b_{n_c}/b'_{n_c}$. Sending $c \to \infty$ leads to $\varepsilon < 0$, which is a contradiction. Thus, we have proved that

$$\limsup_{n\to\infty} \frac{a_n}{a'_n} \leq 1.$$

A symmetric argument proves the converse inequality for lim inf.

Proof of (II) proceeds by applying the already established part (I) to the sequences $\{a_n, n \in \mathbb{N}\}$ and $\{a_n - a'_n, n \in \mathbb{N}\}$ and noting that the relation $b_n \sim b_n - b'_n$ implies $a_n \sim a_n - a'_n$. The proof is complete. □

Using a similar reasoning, one can prove the following.

Theorem 5.2. *Let the triples* (h_1, g, x_0) *and* (h_2, g, x_0) *generate the iterative functions* f_1 *and* f_2*, respectively. Assume that* $\lim_{x\to\infty} f_1(x) = +\infty$*. Then*

I. $h_2(x) \sim h_1(x)$, $x \to \infty$ *implies* $f_2(x) \sim f_1(x)$, $x \to \infty$, *and*
II. $h_2(x) = o(h_1(x))$, $x \to \infty$ *implies* $f_2(x) = o(f_1(x))$, $x \to \infty$.

5.3. **Sufficient condition for the divergence of solutions to (2).** A simple sufficient condition for $\lim_{n\to\infty} a_n = +\infty$ is given as follows.

Proposition 5.1. *Assume that the sequence* $\{a_n, n \in \mathbb{N}\}$ *satisfies (2). If* $I_n \stackrel{P}{\to} \infty$ *and* $\liminf_{n\to\infty} b_n = b > 0$, *then* $\lim_{n\to\infty} a_n = +\infty$.

Proof. From recurrence (2), we obtain

$$\begin{aligned} a_n &= b_n + \sum_{k=1}^{n-1} p_{nk} a_k = b_n + \sum_{k=1}^{M-1} p_{nk} a_k + \sum_{k=M}^{n-1} p_{nk} a_k \\ &\geq b_n + \Big(\inf_{1\leq k<M} a_k\Big) \sum_{k=1}^{M-1} p_{nk} + \Big(\inf_{M\leq k\leq n-1} a_k\Big) \sum_{k=M}^{n-1} p_{nk}. \end{aligned}$$

Sending $n \to \infty$ gives $\liminf_{n\to\infty} a_n \geq b + \inf_{k\geq M} a_k$. Letting $M \to \infty$ leads to $\liminf_{n\to\infty} a_n \geq b + \liminf_{n\to\infty} a_n$, which completes the proof. □

Acknowledgement. I would like to thank my scientific advisor Alexander Iksanov for introducing me into the subject, careful reading of the manuscript, and numerous helpful comments. Thanks are also due to Martin Möhle for drawing our attention to the problem on the asymptotics of the Poisson–Dirichlet coalescent. Finally, I am grateful to the referee for the careful reading of the manuscript.

References

1. Bruhn, V. (1996) Eine Methode zur asymptotischen Behandlungeiner Klasse von Rekursionsgleichungen mit einer Anwendung in der stochastischen Analyse des Quicksort-Algorithmus. *Dissertation, Christian-Albrechts-Universität zu Kiel.*
2. van Cutsem, B. and Ycart, B. (1994). Renewal-type behaviour of absorption times in Markov chains. *Adv. Appl. Prob.* **26,** 988–1005.
3. Delmas, J.-F., Dhersin, J.-S. and Siri-Jegousse, A. (2008). Asymptotic results on the length of coalescent trees. *Ann. Appl. Probab.* **18,** 997–1025.
4. Devillers, O. (1992). Randomization yields simple $O(n \log^* n)$ algorithms for difficult $\Omega(n)$ problems. *Internat. J. Comput. Geom. Appl.* **2**, 621–635.
5. Drmota, M. (2009). *Random Trees: An Interplay between Combinatorics and Probability*, Springer, Berlin.
6. Drmota, M., Iksanov, A., Moehle, M. and Roesler, U. (2007). Asymptotic results concerning the total branch length of the Bolthausen–Sznitman coalescent. *Stoch. Process. Appl.* **117**, 1404–1421.
7. Drmota, M., Iksanov, A., Moehle, M. and Roesler, U. (2009). A limiting distribution for the number of cuts needed to isolate the root of a random recursive tree. *Random Struct. Algorithms* **34**, 319–336.
8. Flajolet, P. and Sedgewick, R. (2008). *Analytic Combinatorics*, Cambridge University Press, Cambridge.
9. Gnedin, A. V. (2004). The Bernoulli sieve. *Bernoulli* **10,** 79–96.
10. Gnedin, A., Iksanov, A. and Möhle, M. (2008). On asymptotics of exchangeable coalescents with multiple collisions. *J. Appl. Probab.* **45,** 1186–1195.
11. Gnedin, A., Iksanov, A., Negadajlov, P. and Roesler, U. (2009). The Bernoulli sieve revisited. *Ann. Appl. Prob.* **19**, 1634–1655.
12. Gnedin, A. and Yakubovich, Y. (2007). On the number of collisions in Λ-coalescents. *Electron. J. Probab.* **12**, 1547–1567.
13. Haas, B. and Miermont, G. (2010+). Self-similar scaling limits of non-increasing Markov chains, Bernoulli, to appear.
14. Hoar, C. R. (1962). Quicksort. *The Computer Journal* **5(1),** 10-16.
15. Greene D.H. and Knuth D.E. (1990). *Mathematics for the analysis of algorithms*, Burkhäuser, Basel.
16. Iksanov, A., Marynych, A. and Möhle, M. (2009). On the number of collisions in beta(2,b)-coalescents. *Bernoulli* **15,** 829–845.
17. Iksanov, A. and Möhle, M. (2008). On the number of jumps of random walks with a barrier. *Adv. Appl. Probab.* **40**, 206–228.
18. Iksanov, A. and Negadajlov, P. (2008). On the number of zero increments of a random walk with a barrier. *Discrete Math. Theor. Comput. Sci.* Proceedings Series Volume AI, 247–254.
19. Karp R.M. (1994). Probabilistic recurrence relations. *Journal of the Association for Computer Machinery* **41,** 1136–1150.
20. Kesten, H. (1980). The number of distinguishable alleles according to the Ohta–Kimura model of neutral mutation. *J. Math. Biol.* **10,** 167–187.
21. Mahmoud H. (1992). *Evolution of random search trees*, Wiley, New York.
22. Meir, A. and Moon J.W. (1974). Cutting down recursive trees. - *Math. Biosci.* **21,** 173–181.
23. Möhle, M. (2010). Asymptotic results for coalescent processes without proper frequencies and applications to the two-parameter Poisson-Dirichlet coalescent. *Stoch. Process. Appl.* **120,** 2159-2173.
24. Neininger, R. and Rüschendorf, L. (2004). On the contraction method with degenerate limit equation. *Ann. Probab.* **32,** 2838–2856.
25. Panholzer, A. (2004). Destruction of recursive trees. In: *Mathematics and Computer Science III*, Birkhäuser, Basel, 267–280.
26. Panholzer, A. (2006). Cutting down very simple trees. *Quaest. Math.* **29,** 211–227.
27. Rösler, U. (1991). A limit theorem for "Quicksort". *RAIRO, Inform. Theor. Appl.* **25,** 85–100.
28. Rösler, U. (2001). On the analysis of stochastic divide and conquer algorithms. *Algorithmica* **29,** 238–261.
29. Sagitov, S. (2003). Convergence to the coalescent with simultaneous multiple merges. *J. Appl. Prob.* **40,** 839–854.
30. Sedgewick, R. (1977). The analysis of Quicksort Programs. *Acta Informatika* **7(4),** 327–355.

Current address: Faculty of Cybernetics, T. Shevchenko National University of Kiev, 01033 Kiev, Ukraine

E-mail address: marynych@unicyb.kiev.ua